A Flexible Tool for Participating, Authoring, and Managing Citizen Science Campaigns On-The-Go

Sunyoung Kim
Human-Computer Interaction Institute, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA
sunyoung.kim@cs.cmu.edu

ABSTRACT
Grassroots participation is a great resource for monitoring environmental phenomena, and collecting and analyzing data in a variety of fields, called Citizen Science. The proliferation of mobile devices has helped to boost citizen science activities. However, the creation of such tools is under the control of developers and accompanying infrastructure that most local organizations, which are often the bodies of citizen science activities, lack. This often hinders citizen science activities from prospering. This work aims to create a visual environment where people without programming skills can build mobile data-collection tools and manage data collectively. The expected contributions include a creation of a tool to prosper citizen science by lowering technical barriers; an understanding of the dynamics of data collection and management through grassroots participation; and implications for designing citizen science tools.

Author Keywords
Citizen science, mobile applications, iPhone, sustainability.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Design

INTRODUCTION
Across HCI, Ubicomp and other relevant domains focusing on citizen science, there have been efforts to shift data collection platforms from traditional paper-and-pencil to mobile devices in parallel with the proliferation of personal mobile devices [4,6,13]. Personal mobile devices are well suited to support citizen science activities as those are carried everywhere and equipped with wireless data transfer, location tracking, and other state-of-the-art technologies. However, technical difficulties in building and distributing the data collection tool sometimes become a barrier in prospering citizen science activities. In this work, we aim to create a visual environment where people without programming skills can easily build and distribute mobile data-collection tools, and manage collected data for their campaigns. We expect that lowering the technical threshold in building data-collection tools will allow small organizations and local groups to collect a mass of necessary data more easily. The activists will be able to spend more time on the campaign itself and the use of collected data, without struggling with building their own tool. Our system will also display the collected data to be shared within volunteers or in public for collective data monitoring and analysis. Additionally, our system will be a repository of existing citizen science activities where organizations announce new campaigns, and where volunteers explore ongoing campaigns to participate in. This will increase accessibility to campaigns, helping small groups and local organizations to recruit participants for the campaigns. In all, we expect that our system will help citizen science activities thrive.

After activating our system, we will analyze its use to understand the dynamics of data collection and management through grassroots participation, in particular but not limited to focusing on the factors that relate to the success or failure of citizen science activities. We anticipate that the patterns of using our system will reflect characteristics of grassroots participation.

Thesis statement
We hypothesize that a simple technical support can help citizen science activities thrive, especially when those activities are simple, micro-level or local events, by:

1. Lowering technical barriers in building and managing mobile data-collection tools
2. Improving the experience in the use of collected data through collaborative monitoring and analysis
3. Increasing accessibility to existing activities by the provision of a repository where volunteers can explore campaigns

To prove our thesis statements and to analyze the use of such system, we built a visual environment where people can: 1) build simple mobile data-collection tools through drag-and-drop interactions, and 2) explore existing

UbiComp' 12, Sep 5 – Sep 8, 2012, Pittsburgh, USA.

campaigns and aggregated data. We will conduct a study of the system's use and impact. To better understand the dynamics of data collection and management through grassroots participation, we will deploy a full-fledged system in the wild.

EXPECTED CONTRIBUTIONS

With the deployment of our system and the analysis of data, we plan to contribute the following:

1. A platform to support prospering citizen science activities
2. An understanding of the dynamics of data collection and management through grassroots participation
3. A set of design recommendations for mobile data-collection tools for citizen science activities

RELATED WORK

Participatory Sensing is an emerging computing paradigm that enables the distributed collection of data by self-selected participants. It allows the increasing number of mobile phone users to share local knowledge acquired by their sensor-equipped devices in diverse domains, e.g., to monitor temperature, pollution level or consumer pricing information [12]. As an example, the CENS group has put great effort toward and been successful in designing various mobile participatory sensing platforms in extensive domains to empower everyday people to systematically measure and share the world [2,6,8]. While many of the participatory sensing systems have been tremendously successful, we explore a more flexible and adaptive citizen science mobile framework that shifts the concept of *participation* away from its typical concept of sensor data collection.

There have been several research efforts to utilize existing systems as an alternative platform by understanding a lack of infrastructure to utilize the prolific mobile devices for collaboration and coordination in participatory sensing. Twitter has been proposed as an existing system that is well suited to facilitate crowd-sourced participatory sensing [5,7]. However, there are several limitations: 1) there is very little flexibility in customizing Twitter, 2) tweets from participatory sensing activities may pollute personal tweet threads, and 3) Twitter is more suitable for alarming time-critical accidents than for collecting observation-type data. There have emerged more appropriate web-based platforms that allow people to collect, share and query sensor data including SensorWeb and Pachube [11,10] Our system mainly focuses on maximizing flexibility in building a mobile citizen science application with any purposes or needs while ensuring mobile interface usability.

Our work focuses on the flexible self-building and wide distribution of mobile citizen science applications. There exists a small collection of emerging tools with such goals such as Ushahidi and ODK [7,9]. While both tools above are powerful in that a user can build a data collection tool with great flexibility. However, they require infrastructure or basic programming skills. Our system requires neither technical skills nor any infrastructure while providing the same, if not more, functionality, flexibility and the ease of use, with the emphasis on the experience of and participating in citizen science activities. EpiCollect is similar to our system in that it allows the easy development of mobile applications without programming skill in conjunction with web applications to collect and manage data [1]. A difference lies in that EpiCollect was developed primarily to help experts (i.e. epidemiologists) collect field data while our system is designed to encompass general use cases in diverse citizen science domains by experts and non-experts alike with a particular focus on an easy distribution of the application.

RESEARCH PROGRESS

Our initial study focused on designing a mobile citizen science application through human-centered design approach and HCI methods. While successful in its goal, this work informed the concept for our second work that aimed to design and analyze a flexible framework with which people without programming skills can build and manage mobile tools for citizen science campaigns. Thus far, we have built the system and conducted case studies with local organization to collect critical feedback about our system in real citizen science settings.

CreekWatch: designing a citizen science application through human-centered design approach

In our first study, we aimed to design a citizen science application that provided scientists with useful data, and a usable interface for untrained volunteers. The purpose of the application was to help collect data using HCI methods to investigate the needs in designing the application. Working with state and local officials, private groups, and volunteers involved in water monitoring, we conducted a series of contextual inquiries to determine what data and format they wanted and the requirements to make the data usable (e.g., format, reliability, standardization), and implemented *CreekWatch*, an iPhone application and website that allow non-expert volunteers to report information about waterways in order to aid water management programs.

This study revealed that the data collected through *CreekWatch* was indeed useful for existing watershed-managing practices and was in immediate use in water and trash management programs as soon as data was reported. While successful in designing a citizen science application, we also noticed that the final output could be appropriate to other citizen science activities without much modification because its interface was simple, and the components for data collection (*e.g.,* a photo, short comment fields, radio buttons, location information etc.) were applicable to and required by a wide variety of citizen science activities. This shed light on the creation of a flexible framework where people without programming skills can build, distribute, and manage mobile tools for citizen science campaigns.

System implementation for a flexible tool

With the findings from the previous study, extensive literature review in the field, and the exploration of existing

citizen science tools, we designed and implemented a system that allows non-experts in programming can build mobile data collection tools. Our system can support a range of citizen science activities and their volunteers. We defined and used several terms in this work: mobile citizen science request for data collection is *a campaign*, a person or organization who builds and manages the campaign is refereed to as *an author*, and an individual who uses the campaign to collect data is *a volunteer*. The system consists of two parts: (1) *a website* where authors build and manage campaigns and where the public can access the list of currently active campaigns along with data visualizations, and (2) *a mobile application* with which volunteers can explore, subscribe, and participate in ongoing campaigns. (See Figure 1)

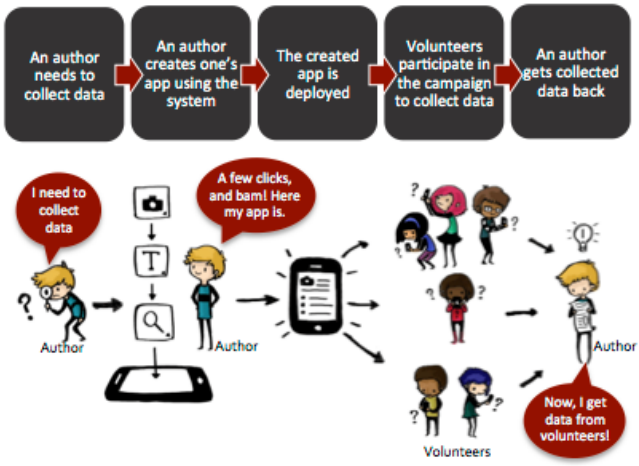

Figure 1. The process of using our system: an author first builds an app → the built app is deployed → volunteers use the app to collect data → collected data is sent back to the author

The main consideration in designing *the website* was that an author without programming skills should be able to easily build and manage a campaign. There are only two steps required in building a campaign: (1) filling out a general information page to describe the campaign and (2) designing a data collection user interface using a simple GUI drag and drop metaphor. Once the campaign is built, the system automatically generates the campaign's hosting webpage where all information and aggregated data for the campaign are displayed. *The mobile application* is used as a tool for volunteers to report collected data. Volunteers can download the application, explore existing campaigns, subscribe ones to participate in, and collect and report data using it. Currently, we are implementing our system, and planning to deploy it in the wild in the near future.

Case studies with local organizations

We presented our system to a range of organizations with ongoing or an interest in initiating citizen science campaigns within their local communities. The domains that the organizations are in action include air quality monitoring, watershed monitoring, and parks conservation. These case studies were designed to provide us critical feedback on the design and use of our system in real citizen science settings. The method used for these case studies is as follows. First, managers in the organizations were gathered to discuss issues encountered in organizing campaigns in terms of volunteer recruitment and data collection. Next, we briefly explained the purpose of our system and presented a working prototype version for them to use and explore. We asked them to browse the system and then to choose one of their own activities and create it within our system. In each community, the manager created a mobile application using the website and tested the resulting app in a separate room. Finally, the community members were gathered together to discuss the possible issues, benefits, and difficulties in using the system. Each community selected one design among those created to use for their activities in the field.

Through the findings from this study, we demonstrated the possibility of our system in strengthening grassroots activities and increasing the public participation in problems to solve directly related to our everyday lives, health, and environmental issues. Overall, our system was extremely well received by all managers participating in the case studies. The most positive comments were around the simplicity in both creating a campaign and reporting data, while two concerns were brought up: quality of data and privacy. For data quality, it was considered important to provide an author with monitoring and authoring features of data both to improve the quality and to ensure the proper use of the data. Compared to the concern regarding data quality, little concern about privacy was raised in the case studies. We assume this is because the target participants in the study are authors who are the recipients of data while privacy concern arouses from the data provider. Additionally, we received several feedbacks about possible positive impacts of our system on citizen science activities in the case study. While different in detail, all fundamentally stem from the same idea: the need to increase public participation. In its nature, citizen science cannot sustain without citizens' participation, and our system can be an additional means to reach a wider population. Using it, any individuals or groups who organize citizen-participatory activities can create mobile applications for activities without any technical resources.

REMAINING RESEARCH AND METHODOLOGICAL APPROACH

The primary goal of this dissertation is to understand the dynamics of citizen science activities with regards to the use of technologies that help citizen science prosper. The main approach is to deploy our system in the wild longitudinally until sufficient activities are undergone so that enough data is aggregated for us to analyze the patterns and dynamics of the use of the system.

The findings from the previous study pointed towards a promising potential for networked digital citizen science for expanding participation and ideation around participatory data collection efforts. However, the previous study has

limitations in that it did not include volunteers in the study who are another crucial stakeholder contributing to data collection and management. As we move forward in the deployment of our system in the wild, we plan to conduct another study to explore the volunteers' use of our system using the same methods applied in the previous study. We will recruit participants among volunteers who regularly collect and report data using our system, and conduct interviews focusing on their experiences. We expect that this study will shows the entire process from start to finish, strengthen our findings from previous study, and/or reveal issues to consider for the success of our system.

After deploying our system in the wild, we will move to the next step of this thesis, understanding the dynamics of citizen science activities with respect to the use of technologies. We anticipate that the nature of networked digital citizen science around participatory data collection efforts will emerge from the longitudinal field deployment of our system. After the system is settled and widely used, we will analyze the patterns and characteristics particularly but not limited to focusing on the factors related to the success or failure of citizen science activities including:

- What types/categories of campaigns benefit from, or are inappropriate for the use of our system?
- What are the characteristics of the interfaces that successful campaigns have in common?
- How and why does each stakeholder (authors/volunteers) utilize our system and what are the benefits and barriers?

We plan to measure followings to prove or disprove each of our thesis statements:

1. Lowering technical barriers – the qualitative evaluation from the field study with authoring stakeholders, and the quantitative analysis of the number of created campaigns
2. Improving the collaborative data monitoring and analysis experience – the qualitative evaluation from the field study with both authoring and volunteering stakeholders
3. Increasing accessibility – the qualitative evaluation from the field study with volunteering stakeholders, and the quantitative analysis of the number of reported data

CONCLUSION

We hypothesized that a simple technical support can help citizen science activities thrive, especially when those are simple, micro-level or local events with little resources in technical skills, by lowering technical barriers in building and managing mobile data-collection tools; improving the experience in the use of collected data through collaborative monitoring and analysis; and increasing accessibility to existing activities by the provision of a repository where volunteers can explore campaigns. Our plan to prove our thesis statements is to deploy the system in the wild and to analyze the use of such system. Our remaining research builds on our preliminary studies that show positive results towards the hypothesis and will ground our work in a real-world context.

BIOGRAPHICAL SKETCH

Sunyoung Kim is a Ph.D. student advised by Dr. Eric Paulos at Carnegie Mellon University's Human-Computer Interaction Institute (HCII). Her research focuses on enhancing everyday health, wellbeing and sustainable life with the use of technology through human-centered design approach. She has published work exploring the use of technologies in domestic environments for everyday health and wellbeing, design space for sustainable everyday life, and enhancing the experience and the success of citizen science activities on mobile platforms. Sunyoung entered the Ph.D. program in the fall of 2008 and expects to complete her dissertation work at the end of 2013.

REFERENCES

1. Aanensen DM, Huntley DM, Feil EJ, al-Own F, Spratt BG, EpiCollect: linking smartphones to web applications for epidemiology, ecology and community data collection. 2009
2. CENS group, http://urban.cens.ucla.edu/
3. E. D. Cristofaro and C. Soriente. Short Paper: PEPSI: Privacy-Enhanced Participatory Sensing Infrastructure. Public Policy, pages 23–28, 2011.
4. Cuff, D., Hansen, M. and J. King (2008), Urban Sensing: Out of the Woods, Communications of the ACM, Volume 51
5. M. Demirbas, M. A. Bayir, C. G. Akcora, Y. S. Yilmaz, H. Ferhatosmanoglu, Crowd-sourced sensing and collaboration using twitter, IEEE International Symposium on A World of Wireless, Mobile and Multimedia Networks, 2010
6. A. Joki, J. Burke, and D. Estrin. Campaignr – a framework for participatory data collection on mobile phones. Technical report, CENS, UCLA, 2007.
7. Y. Liu, Piyawongwisal, P., Handa, S., Liang Yu, Yan Xu, Samuel, A., Going Beyond Citizen Data Collection with Mapster: A Mobile+Cloud Real-Time citizen science experiment, e-Science Workshops, 2011 IEEE 7th International Conference
8. M. Mun, et al., PEIR, the personal environmental impact report, as a platform for participatory sensing systems research. Proceedings of the 7th international conference on Mobile systems, applications, and services - Mobisys '09, page 55, 2009
9. Open Data Toolkit, http://opendatakit.org
10. Pachube, http://pachube.com
11. A. Santanche, S. Nath, J. Liu, B. Priyantha, and F. Zhao. Senseweb: Browsing the physical world in real time. Nashville, TN, April 2006.
12. K. Shilton, J. Burke, D. Estrin, M. Hansen, and M. Srivastava. Participatory Privacy in Urban Sensing. MODUS, 2008.
13. A. Smith, Smartphone Adoption and Usage, July, 2011. http://pewinternet.org/Reports/2011/Smartphones.asp